Public Health
Seattle & King County
HEALTHY PEOPLE. HEALTHY COMMUNITIES.
Alonzo L. Plough, Ph.D., MPH, *Director and Health Officer*

Back-Bridging and Intercensal Estimation Procedures

Jim Allen - October 2002

SYSTEM REQUIREMENTS

- Windows 98+ installed on PC with at least I GB free space
- SPSS for Windows 10+
- Mapinfo 7+
- TGR2MIF, www.GISTools.com, $139.
- Fast internet connection
- BackBridge2000.EXE and BB2000GIS.EXE.

OVERVIEW

The goal of this collection of programs and procedures is to produce, for any selected county in the USA, a file that contains 2000 population data, by 1990 block group geography, for the four major race groups (White, Black, Indian, and Asian) and Hispanic ethnicity. The age groups are indicated by these lower bounds: 0,1,5,10,15,18,20,21,22,25,30,35,40,45,50,55,60,62,65,67,70,75,80,85.

STAGE 1: DOWNLOADING AND BACK-BRIDGING

Start the program **BackBridge2000**. After linking to SPSS, the program will prompt you for the state you wish to work with. It will then download the needed census files (if not already downloaded) from www.census.gov website and store them in c:\BackBridge2000_XX, where XX is the state abbreviation.

You can then select a county in that state. A subfolder of the state folder will be created using the county name, and county records will be extracted from the state files and stored in the county folder. You can then experiment with weights for the four race group White, Black, Indian, and Asian. The weights are used to allocate multiple-race counts to the four major race groups. When you are satified with the weights you wish to use, click Apply to Blocks to create block-specific output files containing Age x Sex x Race data for the four major race groups, plus similar records for Hispanic ethnicity. Output files are FINAL.DBF and FINAL.SAV, the latter being a SPSS system file.

In the output files, age is broken down by the age groups represented in the census data. The lower bounds of the age groups are 0, 5, 10, 15, 18, 20, 21, 22, 25, 30, 35, 40, 45, 50, 55, 60, 62, 65, 67, 70, 75, 80, and 85. The files also contain two numbers which, for each block, are the sex-specific proportions of the 0-4 age group that are under one year old..

STAGE 2: EXTRACT AND CONVERT BOUNDARY FILES

In this stage we prepare the MapInfo boundary files from the downloaded Tiger 2000 files. The MapInfo files will be used in Stage 3.

When stage 1 is complete for a county, the folder Tiger2000 will be a subfolder of the county folder, and will contain a GIS file downloaded from the census website http://www.census.gov/geo/www/tiger/tiger2k. The name of the file is TGRssccc.ZIP, where ss and ccc are the FIPS codes for the state and county.

You must extract all the files from TGRssccc.ZIP into the same folder, Tiger2000.

Then start the program TGR2MIF, which can convert the text files from the census to Mapinfo MIF (MapInfo Interchange Format) files. There are several things to do to get TGR2MIF to do what we need done:

1. Start TGR2MIF.
2. Click Add button, then find the Tiger2000 folder in which you extracted the GIS files (see above). The file TGRssccc.RT1 should become visible in the dialog box. Highlight that file and click Open.
3. A new panel of options will open. Select the "Tiger 2000" version of Tiger (second from bottom), check the "Use this option for all cases" checkbox, then click OK.
4. Another panel with numerous options will open. Select only two options, "Block Groups 90" and "Blocks 2000"., then click OK.
5. The original panel, where you clicked the Add button, will reappear. Click the GO button.
6. Another panel will open. Select the option "Put all outputs in a single directory" and click OK.
7. The program will finish in one to five minutes for counties in WA, and close automatically. You will find MID and MIF files in the "TigerShapes" subfolder of the folder which contained the file that you selected as input in step 2 above.

STAGE 3: COMPUTE SPATIAL PROPORTIONS

Run the program **BB2000GIS,** and select the state and county as you did in stage 1. The program will automate MapInfo and give it the instructions needed to load the MID and MIF files created in stage 2. It will create MapInfo layers for 1990 block groups and 2000 blocks. Output from this stage is the file c:\BackBridge2000_XX\CountyName\tgrssccbgblk.DBF, where XX is the state abbreviation, CountyName is the name of the selected county, and ssccc is the FIPS code for the state and county. The DBF file has fields zBG, zBLK, and propBLKv, where zBG is the 1990 block group, zBLK is the 2000 block code, and propBLKv is the proportion of area in zBLK which falls into zBG.

STAGE 4: COMPUTE 2000 POPULATION FOR 1990 BLOCK GROUPS, 1990 RACE GROUPS

Run the SPSS program "1_2000BGP90.SPS". This program applies the spatial proportions derived in STAGE 3 to the Age x Sex x Race x Block results of back-bridging 2000 census data from STAGE 1. A proportion exists for each intersection of a 1990 block group area with a 2000 block area. After applying the proportions, creating one record for each such intersection, the records are aggregated by 1990 block group to produce the final result, giving 2000 Age x Sex x Race counts by 1990 block groups for a particular county. Output is an SPSS system file named P00BGP20.AG1 containing the fields listed below: That format is consistent with the format used for population files in the Vista desktop system.

Year Place Race T M F
m00 m01 m05 m10 m15 m18 m20 m25 m30 m35
m40 m45 m50 m55 m60 m65 m70 m75 m80 m85
f00 f01 f05 f10 f15 f18 f20 f25 f30 f35
f40 f45 f50 f55 f60 f65 f70 f75 f80 f85.

STAGE 5: COMPUTE INTERCENSAL ESTIMATES BY BLOCK GROUP

Run the SPSS program "2_INTERPOLATE.SPS". This program uses the 1990 population file for a particular county, P90BGP20.AG1, copied from the county's data collection for the Vista desktop system, and the file P00COU20.AG1, produced in STAGE 4 (see above). It produces similarly-formatted files for 1991 through 1999, named $tmp91.sav, $tmp92.sav, etc. Intercensal estimates are computed by simple linear interpolation between the 1990 and 2000 endpoints.

STAGE 6: ADJUST INTERCENSAL ESTIMATES TO OFM AGE X SEX X COUNTY TOTALS.

Run the SPSS program "3_ADJUST.SPS" for each of the nine intercensal files generated in STAGE 5. For a particular year represented by "yy", it uses the file PyyCOU20.AG1 from the Washington State Vista Data Collection. Files should have the creation date 5/14/2002 to ensure that the age x sex totals reflect the estimates produced by OFM to incorporate counts from the 2000 census. For each age x sex cell, the program computes the sum of all age x sex x race x block group values in the file of intercensal estimates, $tmpyy.sav, computed in STAGE 5. Then each of those values is multipled by the ratio x/y, where x is the age x sex cell from the OFM estimates, and y is the sum of the age x sex x race x block group values in that age x sex cell. The output file is named PyyBGP20.AG1, and is compatible with the format of the population files in the Vista desktop system, with the same fields as those listed in the description of STAGE 4 above. Note that the field names identify the sex x age cell (the number in the field name, e.g., "45" in "m45", indicates the lower bound of that field, and implies the upper bound (44) for the preceding field (m40).

Back-Bridging Methodology

Allocation of Multiple-race Counts to Single-Race Categories in Census 2000 Block Groups
Description of Method -- Jim Allen, July 2002

Note -- the descriptions below provide the theoretical framework for the application. The actual implementation combines the Hawiaan race category with the Asian category, and elimates the "Other" race category by (1) ignoring the "Other" component in multiple-race classsifications and (2) by allocating the "Other" single race category to the four main categories (White, Black, Indian, and Asian) proportionally to those four categories after the multiple-race categories have been allocated.

The method for distribution of the multiple-race categories to single-race categories in BackBridge2000 is conceptually the same for two parts of the population, those under 18 and those 18 and over. This happens to be the case because the SF1 files distributed by the US Census Bureau break the multiple-race counts into those two age categories. The explanation below applies to the age < 18 part of the population.

The method is applied to each census block in the state or county selected by the user.

Consider, for each block in the 2000 census, the four age groups 0-4, 5-9, 10-14, and 15-17 (age groups < 18), broken down for males and females, giving eight age x sex cells. Also consider the six single-race-only race categories, here abbreviated as White, Black, Indian, Asian, Hawiaan, and Other. We also use the first letter of each race to denote race components in multiple-race categories. For example, "wba" identifies the multiple-race category "White, Black, and Asian".

The notation and census table references listed below identify the data referred to in this method description. All census tables come from the SF1 release, and thus are complete 100% counts.

T(k) = total population in cell k, for k = 1 to 8. From SF1 table P12.
m(k) = population in two or more race groups in cell k, k = 1 to 8. From table P12g.
s(j,k) = population in race j of cell k, j = 1 to 6, k = 1 to 8. From tables P12a through P12e.
a(j) = the number of people to be distributed to race j from multiple-race <18, table P5 minus P3.

Accumulated allocations from multiple-race counts to

A = a(1) + a(2) + a(3) + a(4) + a(5) + a(6)

We can construct the following table to visualize the data relationships:

	cell 1 M age 0-4	cell 2 M age 5-9	cell 3 M age 10-14	cell 4 M age 15-17	cell 5 F age 0-4	cell 6 F age 5-9	cell 7 F age 10-14	cell 8 F age 15-17	
Total population in cell	T(1)	T(2)	T(3)	T(4)	T(5)	T(6)	T(7)	T(8)	
Pop in mult race group	m(1)	m(2)	m(3)	m(4)	m(5)	m(6)	m(7)	m(8)	
White pop in cell	s(1,1)	s(1,2)	s(1,3)	s(1,4)	s(1,5)	s(1,6)	s(1,7)	s(1,8)	a(1)
Black pop in cell	s(2,1)	s(2,2)	s(2,3)	s(2,4)	s(2,5)	s(2,6)	s(2,7)	s(2,8)	a(2)
Indian pop in cell	s(3,1)	s(3,2)	s(3,3)	s(3,4)	s(3,5)	s(3,6)	s(3,7)	s(3,8)	a(3)
Asian pop in cell	s(4,1)	s(4,2)	s(4,3)	s(4,4)	s(4,5)	s(4,6)	s(4,7)	s(4,8)	a(4)
Hawaiian pop in cell	s(5,1)	s(5,2)	s(5,3)	s(5,4)	s(5,5)	s(5,6)	s(5,7)	s(5,8)	a(5)
Other pop in cell	s(6,1)	s(6,2)	s(6,3)	s(6,4)	s(6,5)	s(6,6)	s(6,7)	s(6,8)	a(6)
									A

After calculating a(j), j = 1 to 6, we use them to distribute each m(k) to the single-race values in cell k. Specifically, the portion of m(k) given to s(j,k) is equal to the proportion that a(j) is of the sum A. Therefore, the distribution to each s(j,k) is computed from the expression m(k) * a(j) / A. After distributing the m(k), the values of s(j,k) in column k sum to the total t(k), so we preserve the age x sex totals for each block group and all block group aggregates, such as tracts and counties.

To calculate the values of a(j), we use tables P3 and P5. P3 contains the population for each multiple race category.
Table P5 contains the population age 18 or more for each multiple race category. To get the population under age 18
for multiple race categories, we subtract each cell in P5 from the corresponding cell in P3. Then, for each multiple
race category, we observe which race components make up the category, divide the associated population into equal shares based on the number of race components, and add one share to each a(j) that correponds to a race
component in the multiple race category.

E Q U A L A L L O C A T I O N E X A M P L E

Category and population	White a(1)	Black a(2)	Indian a(3)	Asian a(4)	Hawiaan a(5)	Other a(6)
wi, 24	12		12			
bao, 15		5		5		5
wbho, 10	2.5	2.5			2.5	2.5
wbiaho, 6	1	1	1	1	1	1

Weighted allocation of the multiple race categories is a simple extension of equal allocation. To simplify the concept somewhat, assume that weights are specified as integers. Then the total number of shares to be distributed is the sum of the weights of the race components present for a particular multiple race category. The shares are then distributed to the a(j) in proportion to those weights. For example, if we assign weights as indicated, the categories and populations in the above table are changed as indicated below.

W E I G H T E D A L L O C A T I O N E X A M P L E

	White	Black	Indian	Asian	Hawiaan	Other
weights =	1	1	2	2	1	1
Category and population	a(1)	a(2)	a(3)	a(4)	a(5)	a(6)
wi, 24	8		16			
bao, 15		3.75		7.5		3.75
wbho, 10	2.5	2.5			2.5	2.5
wbiaho, 6	0.75	0.75	1.5	1.5	0.75	0.75